

# CNN-based Estimation of Abdominal Circumference from Ultrasound images

Jaeseong Jang<sup>\*1</sup>, Ja-Young Kwon<sup>2</sup>, Bukweon Kim<sup>1</sup>, Sung Min Lee<sup>1</sup>, Yejin Park<sup>2</sup>, and Jin Keun Seo<sup>1</sup>

<sup>1</sup>Department of Computational Science and Engineering, Yonsei University, Seoul, Korea

<sup>2</sup>Department of Obstetrics and Gynecology, Institute of Women's Life Medical Science, Yonsei University College of Medicine, Seoul, Korea

## Abstract

The obstetrics and gynecology ultrasound diagnosis is routinely used to check fetal biometry, and due to its time-consuming routine process, there has been great demand of automatic estimation. Automated analysis of ultrasound images is complicated because ultrasound images are patient-specific, operator-dependent, and machine specific. Among fetal biometry, abdominal circumference (AC) is more difficult to make accurate measurement automatically because abdomen has low contrast against surroundings, non-uniform contrast and irregular shape compared to other parameters. This paper proposes a framework for estimation of the fetal AC from 2D ultrasound data by a specially designed convolutional neural network (CNN) which takes account of doctors' decision process, anatomical structure, and the characteristics of ultrasound image. The proposed framework uses CNN to classify ultrasound images (stomach bubble, amniotic fluid, and umbilical vein) and the Hough transform for the measurement of the AC. We tested the proposed method using clinical ultrasound data acquired from 10 pregnant women. Experimental results showed that, with relatively small training samples, the proposed CNN provided sufficient classification results for AC estimation through the Hough transform. This framework showed good performance on most cases and even for ultrasound images deteriorated by shadowing artifacts. However, for oversized fetus cases, when amniotic fluid is not seen or abdominal area was distorted, it could not estimate correct AC.

---

<sup>\*</sup>jaeseongj@yonsei.ac.kr

# 1 Introduction

The ultrasound is most commonly used tool in the obstetrical field for fetal anatomical and functional surveillance. Fetal biometry (estimation of the fetal biparietal diameter (BPD), head circumference(HC), and abdominal circumference(AC)) has been known to be useful to predict intrauterine growth retardation and fetal maturity, and estimating gestational age [1]. Due to its time-consuming routine process, there has been great demand of automatic estimation of these biometry to improve doctor's work flow. Unfortunately, the analysis of ultrasound images is complicated work because ultrasound images are patient-specific, operator-dependent, and machine specific. Hence, automated fetal biometry estimation requires to deal with noisy ultrasound images, which are affected by signal dropouts, artefacts, missing boundaries, attenuation, shadows, and speckle [2].

Among fetal biometry, abdominal circumference(AC), used to monitor fetal growth especially at the last trimester, is harder to measure automatically than head circumference(HC) or femur length(FL) because abdomen has low contrast against surroundings, non-uniform contrast and irregular shape while HC and FL measures bone which has high contrast against surroundings. This paper focuses on this challenging issue of measuring AC automatically.

There have been numerous approaches to extract morphological information from ultrasound data. Most methods are based on the image gradient which requires clear contrast of target from surroundings for a robust result [1, 3, 4, 5, 6]. However due to the low contrast of abdominal boundary, these methods find it hard to guarantee a stable result on AC estimation. Recently, machine learning methods, such as probabilistic boosting tree(PBT)[7], and convolutional neural network (CNN) [8], have been used for fetal biometry measurement. The PBT method is a multi-class discriminative model constructing a tree with its nodes as distinct strong classifiers made by several weak classifiers. By classifying segment structures in ultrasound images, this method estimates fetal biometry parameters [7]. Although this approach showed some notable results, it required complex well annotated data to train the tree. The CNN method, which shows great successes in the object recognition recently, was also applied in fetal biometry by analysing high level features from the ultrasound image data. This method in clinical environment faced obstacles: (i) It is difficult in collecting sufficient data for training. (ii) It is difficult to cope with serious artifacts including shadowing artifacts [8].

In the proposed method, we tried to increase classification performance with insufficient data and deal with the artifacts by contain ultrasound propagation direction and depth as well as multiple scale patches as inputs. The method classifies images patches from a ultrasound image into anatomical structures so that the classification allows to verify acceptability of given abdominal plane. By detecting anatomical structures in a fetal abdomen, the center of the fitting ellipse can be properly initialized in a semantically segmented image, which leads to the fully automatic AC measurement in a given ultrasound image. Followed by initialization, we estimate AC of the accepted plane by using the randomized

Hough transform. We validated our method using an ultrasound data of the AC measurement from fetuses at 20-34 weeks of gestation. 3 trained clinicians evaluated accepted abdominal planes and estimated AC by the method.

The major contribution of our work are the followings:

- We develop a specialized CNN structure which takes account of sonographers' decision process by considering characteristics of the ultrasound imaging. The proposed CNN structure shows high training performance in spite of relatively less training samples.
- We develop a framework which combines the CNN and the randomized Hough transform, to complement each other. The CNN provides the evidence for AC plane evaluation and pre-processing of a ultrasound image for AC estimation, simultaneously. By the combination, we can achieve a stable AC estimation than a mathematical model is used alone.

## 2 Methods

This section proposes a method which combines a CNN, Hough transform and a method for final determination. We begin with explaining the characteristics anatomical structure found in the axial image of the fetal abdomen for AC measurement. As described in Figure 1, the anatomical landmarks on a true axial plane of the abdomen are stomach, umbilical vein, portal sinus, vertebral body, and rib. From this facts, we first constructed a CNN which segments the

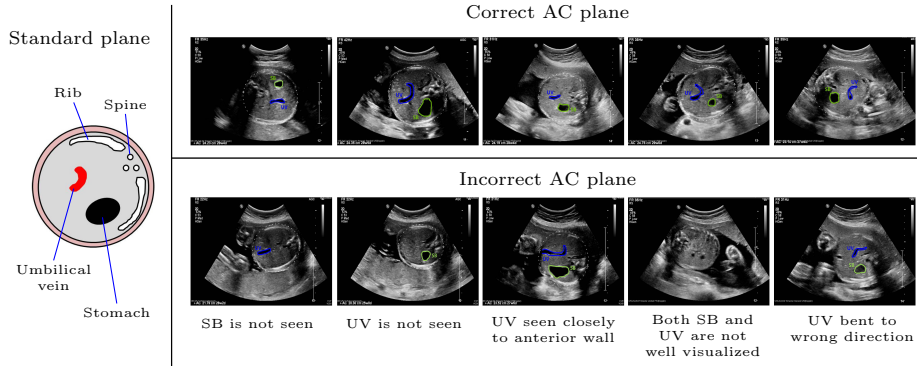


Figure 1: Fetal abdominal ultrasound images and anatomical structures. In standardized abdominal ultrasound images, UV and SB are observed and UV bends against SB.

important fetal abdominal anatomical structures. Using this segmented image, we specified region of interest (ROI) by Hough transform. With this ROI we assessed whether the plane is acceptable or not. If accepted, we measured AC with this ROI by placing the measurement callipers on the skin surface.

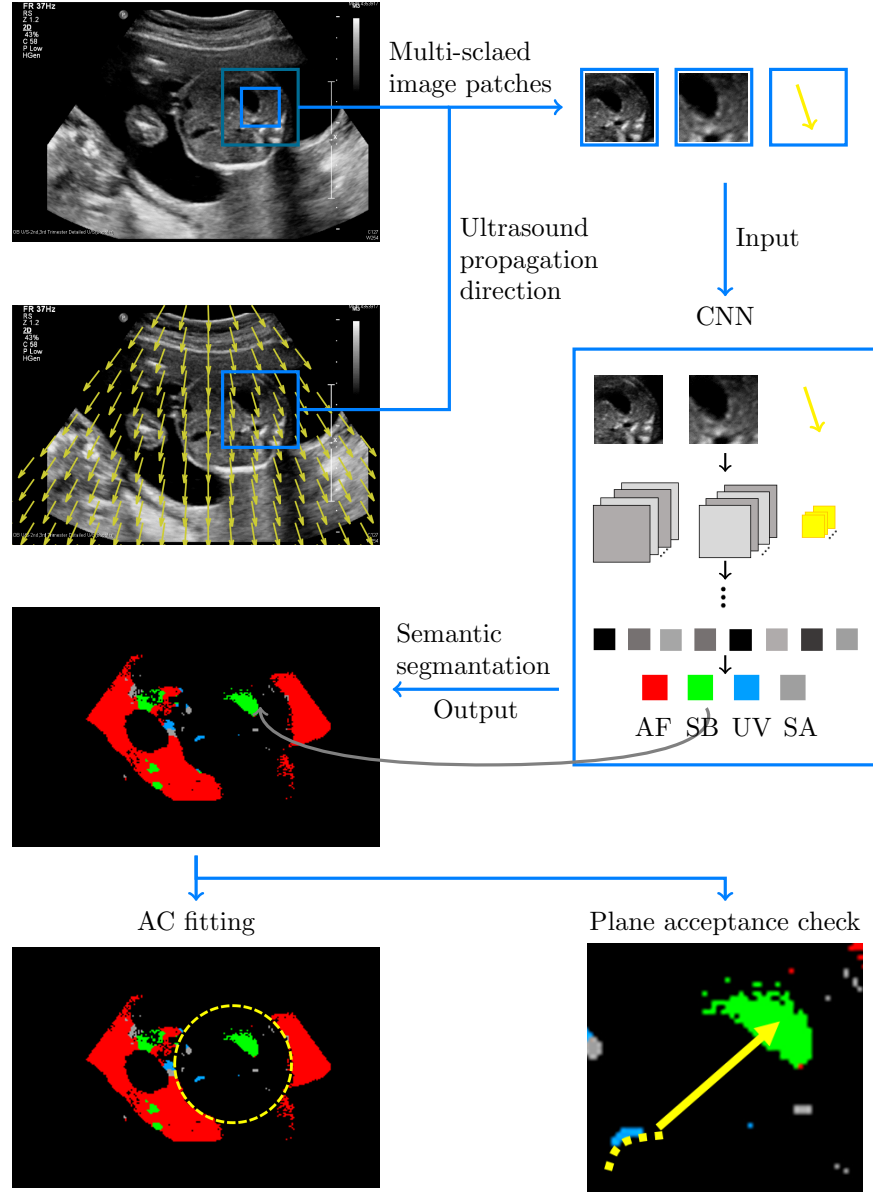


Figure 2: Overall process of the proposed framework. The proposed framework performs semantic segmentation by using a CNN, AC measurement, and plane acceptance check.

## 2.1 Convolutional Neural Network

CNN is a type of artificial neural networks inspired by visual information processing in the brain. To recognize complex features from the visual information, CNN consists of several layers which extract and repeatedly combine low-level features for the composition of high-level features. The composed high-level features are used for CNN to classify an input image.

For the process and the simplicity of computation, many CNN structures are described by combinations of convolutional, pooling, and fully-connected layers. Convolutional layer (C-layer) extracts higher-level features by convolving received feature maps from the previous layer and activating the convolved features. To be precise, suppose that  $j$ -th layer has  $N_j$  nodes (neurons) whose output feature map is  $\mathbf{h}_i^j$ . Then,

$$\mathbf{h}_i^j = g \left( \sum_{n=1}^{N_{j-1}} \mathbf{W}_{in} * \mathbf{h}_n^{j-1} + b_i^j \right) \quad \text{for } j = 1, \dots, N_j$$

where  $\mathbf{W}_{in}$  is the convolutional filter connecting  $i$ -th node (neuron) on the  $j$ -th layer and  $n$ -th node (neuron) on the  $j-1$ -th layer,  $b_i^j$  is the bias for the  $i$ -th output on the  $j$ -th layer, and  $g$  is the activation function. In this article, rectified linear unit (ReLU)  $g(x) = \max(0, x)$  is used as the activation function.

A C-layer usually is followed by a pooling layer (P-layer) which reduces the dimension of feature maps by "max pooling". The max pooling downsamples the input feature maps by striding a rectangular receptive field and taking the maximum in the field. For example, we use  $2 \times 2$  receptive field with stride 2 in our research, which reduces the dimension of feature map by 1/4.

After couples of pairs of C-layer and P-layer, a fully-connected layer (F-layer) integrates high-level features and produces compact feature vectors:

$$\mathbf{h}_i^j = g \left( \sum_{n=1}^{N_{j-1}} \langle \mathbf{W}_{in}, \mathbf{h}_n^{j-1} \rangle + b_i^j \right) \quad \text{for } j = 1, \dots, N_j$$

for the  $j$ -th F-layer. Note that we use inner-product rather than convolution between the filter  $\mathbf{W}_{in}$  and the input feature map  $\mathbf{h}_n^{j-1}$ . Like the C-layers, ReLU is used as the activation function of the F-layers in our research. On the final layer, say  $J$ -th layer, the output layer produces the posterior probability for each class by the softmax function:

$$\text{softmax}(\mathbf{o})_i = \frac{e^{o_i}}{\sum_{n=1}^N e^{o_n}}$$

where the output  $\mathbf{o} = [\sum_{n=1}^{N_{J-1}} \langle \mathbf{W}_{in}, \mathbf{h}_n^{J-1} \rangle + b_i^J]_{i=1, \dots, N_J} = [o_1, \dots, o_{N_J}]$ .

## 2.2 Proposed CNN Structure

As mentioned in the background, characteristics of classification process in ultrasound images should be applied to a CNN structure for the completeness of

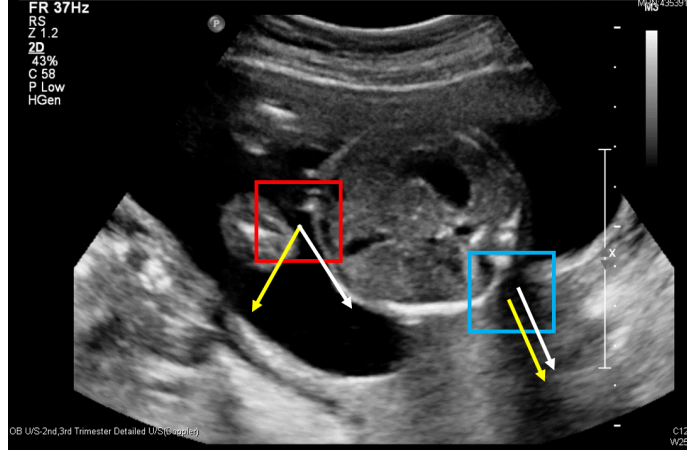


Figure 3: The relation between ultrasound propagation direction (white arrows) and image pattern direction (yellow arrows) in image patches. In the image patch corresponding to shadowing artifact (light blue box), the image pattern is strongly related to the ultrasound propagation direction compared to the patch corresponding to AF (red box).

the structure. We begin with the observations of additional information of ultrasound images integrated with local pattern to determine a given local pattern as followings:

1. We conclude a image pattern as the shadowing artifact not only by its image pattern but also by expected ultrasound propagation direction and the position of hard material (spine, ribs).
2. According to the position of patches, their speckle patterns vary with position in spite of same organ due to the characteristics of point spread function of ultrasound imaging.
3. A proper size of an image patch may vary with anatomical structures which we want to detect.

Based on these observations, we use multiple image patches of 2 sizes (approximately, 10% and 20% of image height) with expected ultrasound propagation direction with respect to the ultrasound probe at the patch position. For example, the propagation direction  $(u, v)$  can be simply modelled by

$$u = \frac{x}{\sqrt{x^2 + y^2}}, v = \frac{y}{\sqrt{x^2 + y^2}}$$

where  $(x, y)$  is the representation of the patch position with respect to the probe. The output of the proposed CNN structure is  $1 \times 1 \times 4$  vector which corresponds to 4 categories which include the shadowing artifact (SA) and the main

Input	Small image		
	$27 \times 27 \times 1$		
	Type	Maps	Filter size
	C	$24 \times 24 \times 25$	$4 \times 4$
	P	$12 \times 12 \times 25$	$2 \times 2$
	C	$8 \times 8 \times 25$	$5 \times 5$
	P	$4 \times 4 \times 25$	$2 \times 2$
Output	F	$1 \times 1 \times 8$	-

Table 1: The proposed CNN structure for plane acceptance check. The output of the network has 8 classes which correspond to the 8 directions.

anatomical structures in the standardized abdominal plane: stomach bubble (SB), umbilical vein (UV), amniotic fluid (AF). The reason why the shadowing artifact is included as one class is to prevent from classifying shadowing artifact *in the fetal region* as amniotic fluid, which could cause a serious error for the fetal AC estimation. According to the 4 classes, image patches sampled at dark point ( $\leq 0.05 \times \text{Max Intensity}$ ) in a given ultrasound image are classified during the semantic segmentation step. The proposed CNN consists of three branches of successive layers which separately handle image patches of multiple scales and propagation direction after the input layer which separates an input data into the three types of data. The separated data passes through the corresponding branch to extract desired image features and analyze propagation direction.

As described in Table 4, two branches for image analysis basically consist of pairs of convolutional and max-pooling layer, and a full-connected layer. For the first and second convolutional layers in the branch for small image analysis,  $11 \times 11$  and  $5 \times 5$  filters were used, respectively, while  $11 \times 11$  and  $5 \times 5$  filters were used in the branch for large image analysis. And  $2 \times 2$  max-pooling was used with 2 the stride step. Ultrasound propagation direction is analyzed through one fully-connected layer to detect the propagation direction. Such feature vectors produce by full-connected layers in the three branches were concatenated into one feature vector for the decision. The concatenated feature vector passes through two full-connected layers and the CNN classifies a given data into 4 classes.

### 2.3 Plane Acceptance Check

In this section, we evaluate suitability of selected plane to determine whether the plane is acceptable. Firstly, we checked the presence of SB and UV in the segmented image from previous step. If either SB or UV was absent in the image, the selected plane was not accepted for the AC measurement. After checking the presence of SB and UV, we assessed the anatomic configuration of UV and SB by estimating a bending direction of UV and comparing the direction to the position of SB. Based on the guideline in Figure 1, UV appears to be bent against SB as described in the standardized fetal abdominal plane. The bending

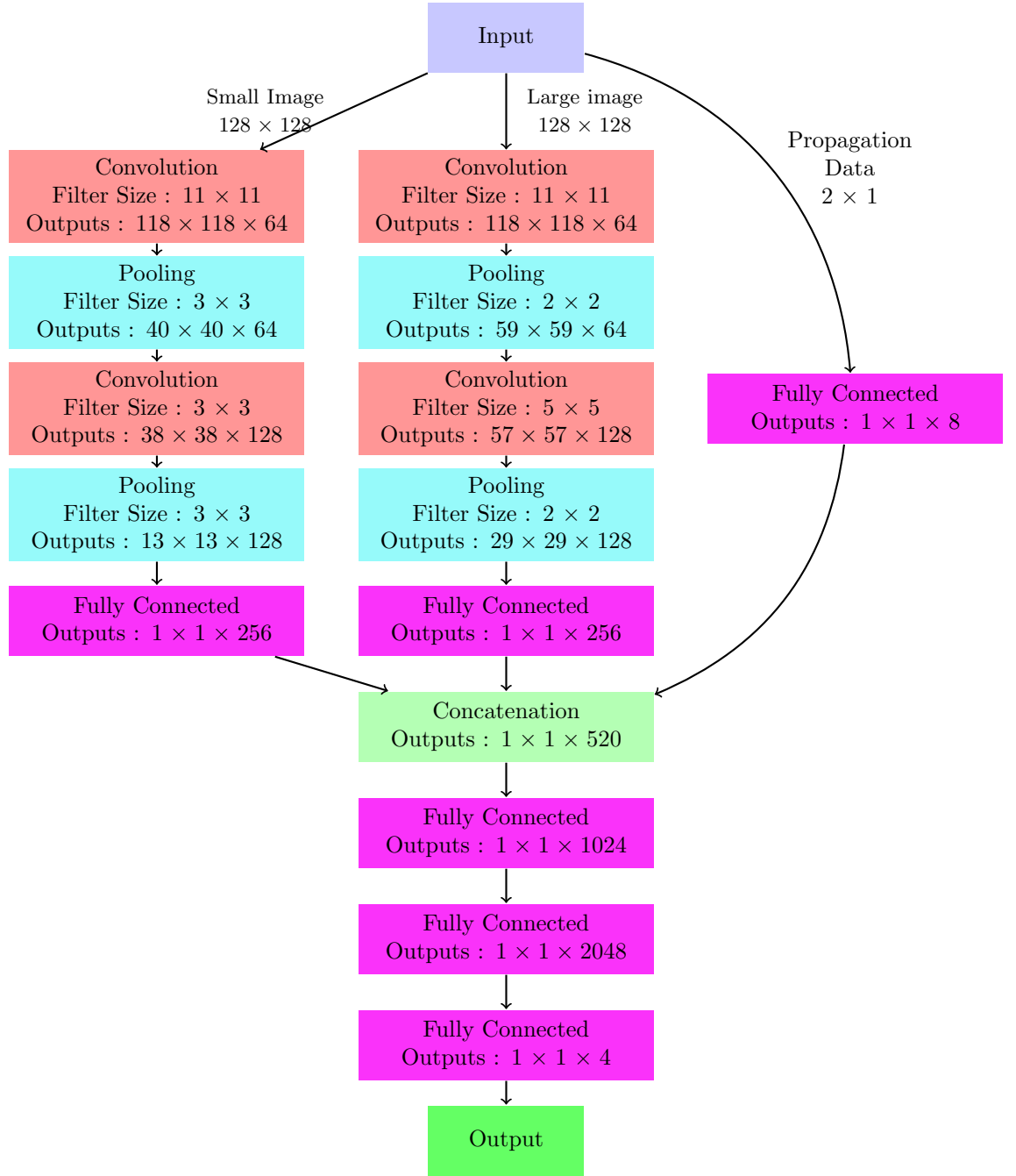


Figure 4: The proposed CNN structure. The CNN uses input data as a combination of image patches of multiple sizes and ultrasound propagation direction. From the image patches and the propagation directions, feature vectors are extracted and combined to classify a given image patch.



direction of UV was detected by using a CNN which classifies a given shape into straight or bending shape against 8 directions. If a given UV is classified as a straight shape, the given frame was not accepted. Finally, the final score was evaluated by taking inner-product between the direction and position vector of SB with respect to UV, which is defined by the vector between the center of mass of SB and that of UV. Then the selected plane and AC estimation were accepted if the score was more than 0.2. Fetal transabdominal plane was obtained by 2D ultrasound examinations performed with an IU22 (Phillips, Seoul, Korea) ultrasound machine by an operator using 2–6-MHz transabdominal transducer. Ultrasound images were reviewed by each of the 3 ultrasound experts including the operator and scored 1 point if the image was correct AC measurement plane, or 0 if non-acceptable according to the criteria described above. The image with total score of 2 or more was considered acceptable. Simultaneously, the proposed CNN in the previous section was and the proposed anatomic configuration model were used to evaluate the suitability of selected plane.

## 2.4 Measurement agreement

The ellipse detection method based on Hough transform[9] was applied to the semantic segmentation image for quantifying AC by assuming that the fetal abdomen has an ellipsoidal shape modelled by the following form with five parameters  $a$ ,  $b$ ,  $c$ ,  $p$ , and  $q$ :

$$a(x - p)^2 + 2b(x - p)(y - q) + c(y - q)^2 = 1 \quad (1)$$

where  $ac - b^2 > 0$ . By initializing the center of the ellipsoidal shape based on the detected SB and UV regions, we found the parameters which has the highest confidence to represent the shape. Thirty transabdominal images demonstrating

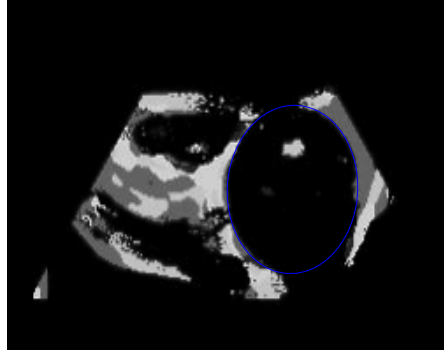


Figure 5: Detection of fetal abdominal contour using ellipse detection method

proper landmarks for true axial plane for AC measurement were obtained by an expert. The AC measurement was performed either manually by other experts or by the proposed method. The assessment of interobserver variability between

two experts and AC values comparison between manual and CNN was performed.

### 3 Results

#### 3.1 Training and test data set

For training, we used 56 cases of fetal abdominal ultrasound images provided by the department of obstetrics and gynecology at the college of medicine of Yonsei university. From the provided images, 13261 pairs of multiple scale image patches were extracted with ultrasound propagation direction in those patches. To augment the training dataset, the dataset were mirrored only in the horizontal direction with the sign of the x-direction of propagation changed. Mirroring the dataset in the vertical direction was not considered because it is meaningless to train images mirrored in vertical direction because ultrasound propagates from the top to bottom. *Caffe* was used to implement and train the proposed CNN [10].

#### 3.2 Semantic segmentation

As described in Figure 6, we performed semantic segmentation for pixels which has low intensity. The segmentation results are represented color maps which have red, green, blue, and gray color corresponding to SB, UV, amniotic fluid, and shadowing artifact. When relatively high intensity appears in SB region (the first case in Figure 6), the proposed framework fails to detect the region. We can observe false SB and UV regions in AF or SA regions (the second and third cases in Figure 6). In the SB and UV region of fetuses, false AF and SA regions appear, too, and a portion of SB region connected with AF is classified as AF region (the forth case in Figure 6).

#### 3.3 Acceptance Check

We trained a CNN for the plane acceptance check by using the semantic segmentation images made from the previous step. In the transabdominal planes accepted by the 3 ultrasound experts, the bending direction is defined as the vector between UV and SB regions. In Figure 7, we plot accepted check results for selected fetal transabdominal ultrasound images. with anatomic In the third results, the transabdominal plane is accepted by the experts, but our framework did not acceptable plane. That's because the semantic segmentation is applied only to the dark region in the previous step, which results underestimation of UV region.

#### 3.4 AC Estimation

With accepted images, the circumference of a detected ellipsoidal shape was estimated by using the parameters in (1). For selected ultrasound images, ab-

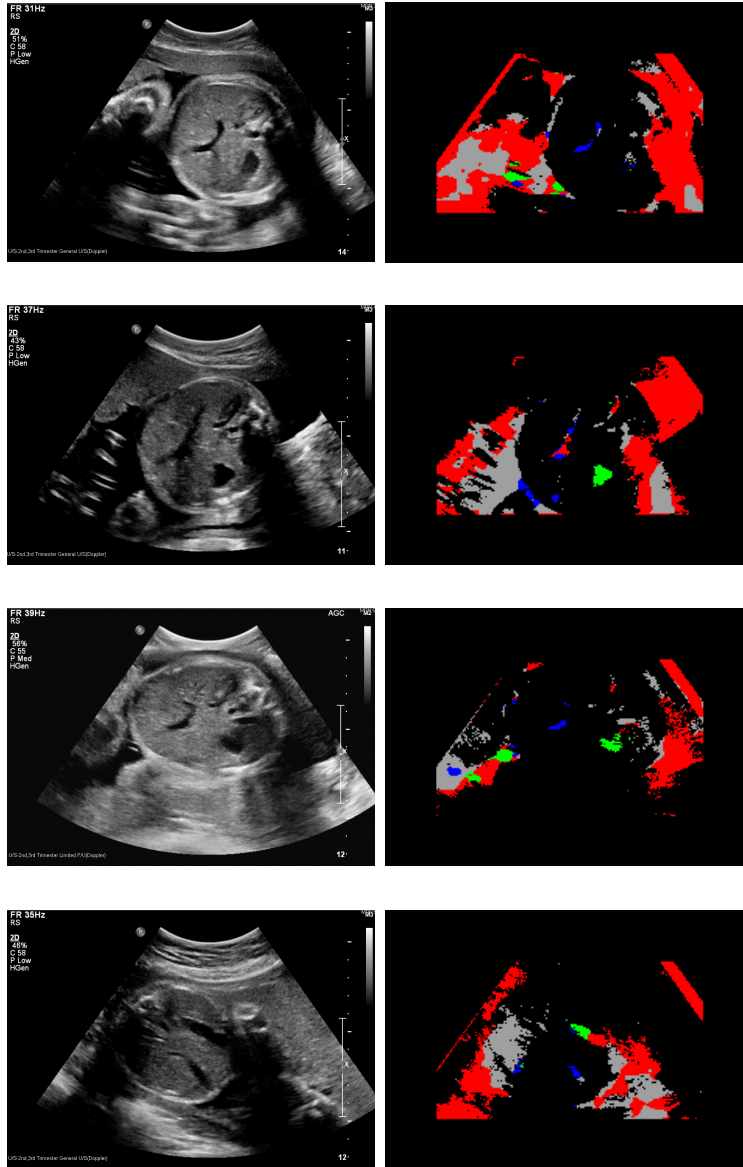


Figure 6: Original ultrasound images with marks for anatomical structures (left column) and corresponding semantic segmentation results obtained by the proposed CNN structure (right column). Red, green, blue, and gray regions correspond to AF, SB, UV, and AF regions, respectively.

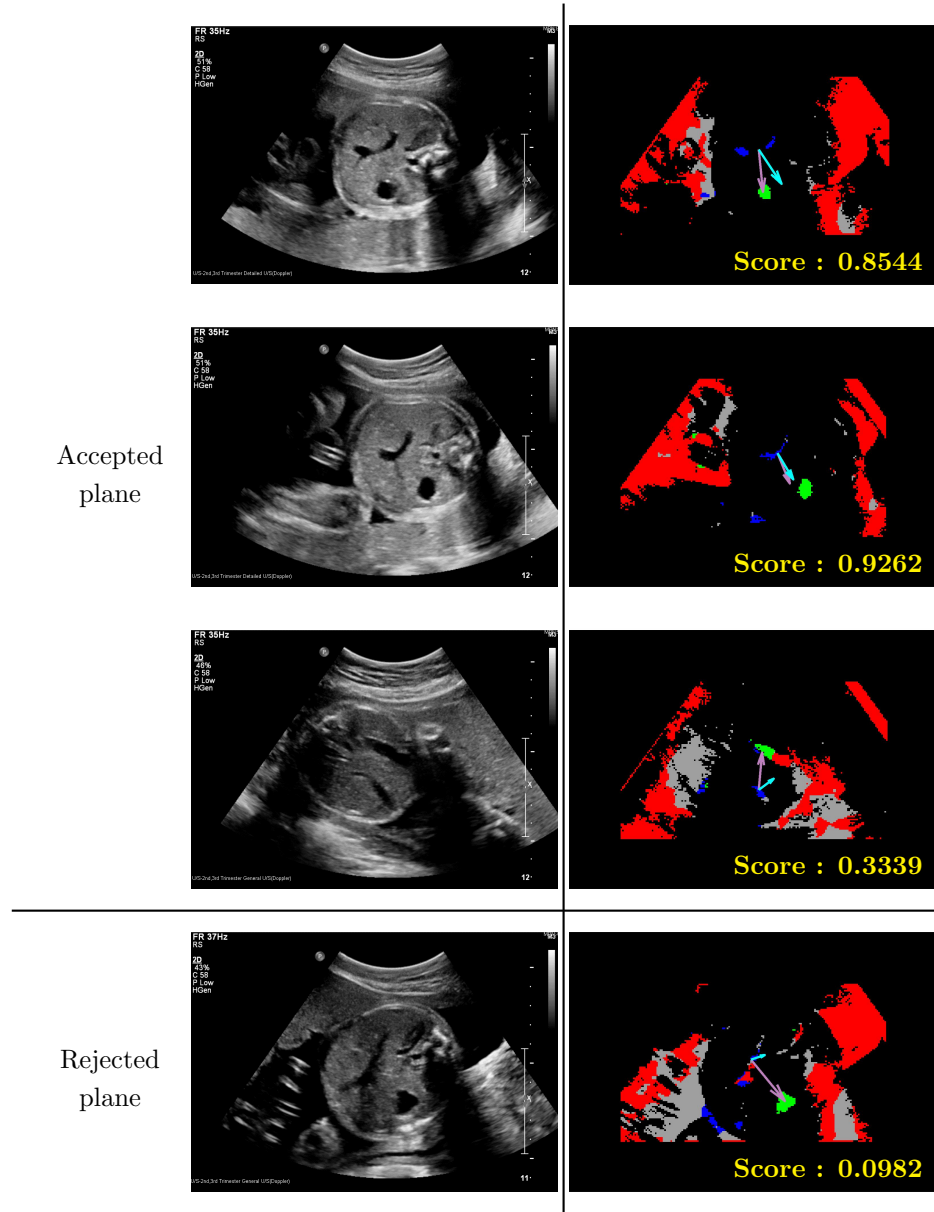


Figure 7: Performance evaluation for discrimination of appropriate AC plane. Purple arrow represents stomach position from picked vein and cyan arrow represents calculated bent direction of the vein

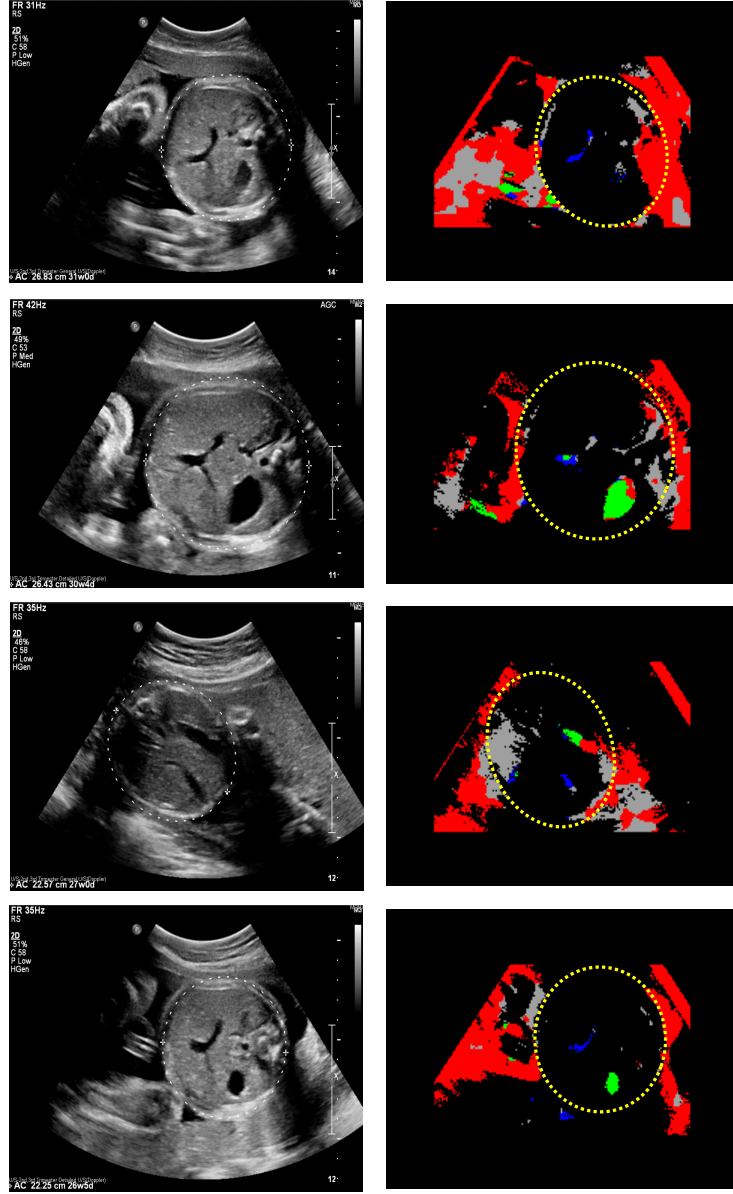


Figure 8: Original ultrasound images with AC contours marked by clinicians (left column) and corresponding AC tracking results by the proposed framework (right column). Tracking results are represented as a yellow contour in each semantic segmentation result.

dominal contours selected by the experts and ellipsoidal contour are compared in Figure 8. We compared the AC estimated by the proposed framework with the AC measurement performed by the experts (Figure 8).

## 4 Discussion

Although CNN showed good performance in image recognition recently, CNN demands to collect a large amount of training data in order to achieve satisfactory pixel-wise classification results. Unfortunately, due to the limitation of gathering clinical data, it is difficult to collect enough data to guarantee satisfactory classification for various cases of ultrasound images. If a CNN is trained only with image data due to lack of its physical characteristics, the number of our training data required become larger. We tried to evade this problem by designing modality-specific structured CNN and got notable improvement in training performance, however there still remained large possibility of improvement with more data. Admitting this limitation, we propose that machine learning in medical field might find breakthrough by focusing on developing modality-specific structure of CNN for more accurate application.

We did only segmentation with CNN because there already existed many fine methods measuring elliptic object in segmented images. So, for using many existing methods, CNN may replace some edge detecting/enhancing, or denoising mathematical models which are required for many conventional methods. In short, CNN paired with well developed other methods may work better than standing alone.

Our method extracts abdominal contour from amniotic fluid. It makes difficult to extract the contour from an oversized fetus because they lack amniotic fluid. And the fact that oversized fetus sometimes have distorted boundary makes it even harder to measure AC. In order to overcome these problems, detecting fetal ribs and placenta of mother might be required.

For further application, we need to develop the system to determine a proper AC plane given whole video frames or 3D volume data. Since our method meets most of the essential parts of full automation, we believe this method can accomplish the goal with certain amount of improvements such as a smart slice selection in 3D volume.

## 5 Conclusion

In this paper, we proposed a method which automatically estimate AC from ultrasound image and evaluate the reliability fetal abdominal ultrasound image. This method shows good performance with limited data on most cases even for ultrasound images deteriorated by shadowing artifacts. It does show some limitations on oversized fetus cases, however, we believe that this is promising way to approach fully automatizing the whole procedures for fetal biometry parameter in the future.

## References

- [1] Vikram Chalana, Thomas C Winter, Dale R Cyr, David R Haynor, and Yongmin Kim, “Automatic fetal head measurements from sonographic images,” *Academic Radiology*, vol. 3, no. 8, pp. 628–635, 1996.
- [2] Sylvia Rueda, Sana Fathima, Caroline L Knight, Mohammad Yaqub, Aris T Papageorgiou, Bahbib Rahmatullah, Alessandro Foi, Matteo Maggioni, Antonietta Pepe, Jussi Tohka, et al., “Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: a grand challenge,” *IEEE Transactions on medical imaging*, vol. 33, no. 4, pp. 797–813, 2014.
- [3] Sayan D Pathak, DR Haynor, and Y Kim, “Edge-guided boundary delineation in prostate ultrasound images,” *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1211–1219, 2000.
- [4] Yongjian Yu and Scott T Acton, “Speckle reducing anisotropic diffusion,” *IEEE Transactions on image processing*, vol. 11, no. 11, pp. 1260–1270, 2002.
- [5] Sandra MGVB Jardim and Mário AT Figueiredo, “Segmentation of fetal ultrasound images,” *Ultrasound in medicine & biology*, vol. 31, no. 2, pp. 243–250, 2005.
- [6] Weiming Wang, Jing Qin, Lei Zhu, Dong Ni, Yim-Pan Chui, and Pheng-Ann Heng, “Detection and measurement of fetal abdominal contour in ultrasound images via local phase information and iterative randomized hough transform,” *Bio-medical materials and engineering*, vol. 24, no. 1, pp. 1261–1267, 2014.
- [7] Gustavo Carneiro, Bogdan Georgescu, Sara Good, and Dorin Comaniciu, “Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree,” *IEEE transactions on medical imaging*, vol. 27, no. 9, pp. 1342–1355, 2008.
- [8] Hao Chen, Dong Ni, Jing Qin, Shengli Li, Xin Yang, Tianfu Wang, and Pheng Ann Heng, “Standard plane localization in fetal ultrasound via domain transferred deep neural networks,” *IEEE journal of biomedical and health informatics*, vol. 19, no. 5, pp. 1627–1636, 2015.
- [9] Yonghong Xie and Qiang Ji, “A new efficient ellipse detection method,” in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. IEEE, 2002, vol. 2, pp. 957–960.
- [10] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell, “Caffe: Convolutional architecture for fast feature embedding,” *arXiv preprint arXiv:1408.5093*, 2014.